DIMACS Technical Report 98-46

October 1998

On Crossing Sets, Disjoint Sets and the Pagenumber [1]

by

Farhad Shahrokhi and Weiping Shi

Department of Computer Science

University of North Texas

Denton, Texas 76203

[1]Research of the first author was supported by NSF grant CCR-9528228. A part of this work was done when the first author was visiting DIMACS in Spring of 1998. Research of the second author was supported by NSF grant MIP-9309120. A portion of this paper was presented in COCOON96.

DIMACS is a partnership of Rutgers University, Princeton University, AT&T Labs-Research, Bell Labs, Bellcore and NEC Research Institute.

DIMACS is an NSF Science and Technology Center, funded under contract STC–91–19999; and also receives support from the New Jersey Commission on Science and Technology.

ABSTRACT

Let $G = (V, E)$ be a t-partite graph with n vertices and m edges, where the partite sets are given. We present an $O(n^2 m^{1.5})$ time algorithm to construct drawings of G in the plane so that the size of the largest set of pairwise crossing edges, and at the same time, the size of the largest set of disjoint (pairwise non-crossing) edges are $O(\sqrt{t \cdot m})$. As an application we embed G in a book of $O(\sqrt{t \cdot m})$ pages, in $O(n^2 m^{1.5})$ time, resolving an open question for the pagenumber problem. A similar result is obtained for the dual of the pagenumber or the queuenumber. Our algorithms are obtained by derandomizing a probabilistic proof.

1 Introduction and Summary

1.1 Preliminaries

Throughout this paper $G = (V, E)$ is an undirected graph with $|V| = n$ and $|E| = m$. A *linear ordering* of a set S is a bijection from S to $\{1, 2, \ldots, |S|\}$. Let h be a linear ordering of V. Consider a drawing of G that is obtained by placing the vertices along a straight line in the plane in the order specified by h, and drawing each edge ij above the line between $h(i)$ and $h(j)$. We assume that an edge does not cross itself, any two edges cross at most once, and no two edges touch each other. Moreover, we assume that only non-incident edges cross and no three edges cross at the same point. In sequel we refer to such a drawing as a *linear drawing* of G, and denote it by (G, h).

A matching is a set of independent edges in G. Let M, $|M| \geq 1$, be a matching in G, and h be a linear ordering of V. M is a *crossing set* if every pair of edges in M cross in (G, h). M is a *disjoint set* if all edges in M are pairwise non-crossing in (G, h). (Note that, when $|M| = 1$, that is, when M contains only one edge, we refer to it as a crossing set.) Note that a crossing set and a disjoint set in (G, h) correspond to a clique and an independent set of vertices, respectively, in the intersection graph of the drawing. (Here, the vertices of the intersection graph are the edges of G, and every edge in the intersection graph corresponds to a pair of crossing edges in (G, h).)

Crossing and disjoint sets are well studied problems in combinatorial geometry [16]. For instance, the minimum number of edges that G must have in order for every linear drawing to contain at least one disjoint or crossing set of size k, has been investigated. See Pach and Töröcsik [18] for disjoint sets, and Capoyleas and Pach for crossing sets [5]. Similar problems have been extensively studied in a more general setting when the underlying drawing is an arbitrary drawing of G in the plane [2, 17, 1, 25]. For any linear drawing, the lower bounds of $\Omega(\frac{m}{n})$ and $\Omega(\sqrt{\frac{m}{n}})$ on the sizes of the largest crossing and disjoint sets, respectively, easily follow from the results in [5] and [18]. A natural problem motivated by these results is to determine in polynomial time, a linear ordering h so that the sizes of the largest crossing set and the largest disjoint set in (G, h) are small. Thus, for the resulting intersection graph the sizes of the cliques, and the sizes of the independent set of vertices corresponding to the disjoint sets, are small.

A *book* consists of a *spine* and some number of *pages*. The spine is a line and each page is a half-plane with the spine as its boundary. A *book embedding* of $G = (V, E)$ consists of placing the vertices of G on the spine in an order specified by a linear ordering of V, and an assignment of edges to pages so that edges assigned to the same page can be drawn on that page without crossing each other. The *pagenumber*, or *stacknumber*, of G is the smallest k so that G can be embedded in a k-page book [4, 6, 20, 21, 24]. One may re-state the pagenumber problem as the determination of a linear ordering h so that the intersection graph of (G, h) has the smallest possible chromatic number. Heath, Leighton, and Rosenberg [11], and Heath and Rosenberg [12] studied a dual problem to the pagenumber and called it the *queuenumber*. Their notion of the queuenumber of a layout is equivalent to the size of the largest nested disjoint set in the linear drawing (G, h), for a given linear ordering h.

Computing the pagenumber is NP-complete even if the linear ordering is fixed [8]. Chung et. al [6] showed that the pagenumber of any d-regular graph is $O(d\sqrt{n})$ through a Las Vegas algorithm. Malitz [13] improved the upper bound to $O(\sqrt{m})$, using his Las Vegas algorithm. One of the most important results in this area is due to Heath and Istrail [10] who showed that the pagenumber is $O(g)$, where g is the (orientable or non-orientable) genus of G. Given an initial embedding of G on a surface, they presented a linear time deterministic algorithm to embed G in a book whose number of pages is linear in the genus of the surface. Using the framework in [10], Malitz [14] improved this result to $O(\sqrt{g})$ employing a Las Vegas algorithm, provided that an initial embedding is given. The algorithms in [10] and [14] require an embedding of G on a surface of genus g which in general is an NP-hard problem.

Finding a deterministic polynomial time algorithm for embedding a graph G in $O(\sqrt{m})$ pages has been one of the most important problems in this area and remains open even for degree bounded graphs. Concerning the dual of the pagenumber, Heath and Rosenberg [12] noted the existence of a layout whose queuenumber is $O(\sqrt{m})$, but the construction of this layout in deterministic polynomial time remains open.

1.2 Our Results

Let (G, h) be a linear drawing of G. We denote by $C(G, h)$ and $I(G, h)$, the size of the largest crossing and disjoint sets, respectively, in (G, h). Our studies of the page number problem are heavily focused on the properties of the crossing sets. Our main result is to show that:

Theorem 1.1 *Let $G = (V_1, V_2, E)$ be a bipartite graph on n vertices and m edges. There is an $O(n^2m^{1.5})$ time algorithm that finds a linear drawing (G, h), so that $C(G, h) \leq \mathbf{e}\sqrt{m}$, and $I(G, h) \leq \mathbf{e}\sqrt{m}$, where $\mathbf{e}$ is the base of the natural logarithm. Moreover, G can be embedded in a book of $C(G, h) \leq \mathbf{e}\sqrt{m}$ pages in $O(n^2m^{1.5})$ time.*

The proof technique in Theorem 1.1 employs the method of conditional probabilities for derandomizing a probabilistic proof. For more on the probabilistic method, see Alon, Spencer and Erdős[3], and also Raghavan [19]. Moreover, we easily extend the above result from bipartite graphs to arbitrary graphs in the following form:

Theorem 1.2 *Let $G = (V, E)$ be a t-partite graph on n vertices and m edges, where the partite sets are given. There is an $O(n^2m^{1.5})$ time algorithm that finds a linear drawing (G, h) so that $C(G, h) \leq \mathbf{e}\sqrt{(t-1) \cdot m}$ and $I(G, h) \leq \mathbf{e}\sqrt{(t-1) \cdot m}$. Moreover, G can be embedded in a book of $\mathbf{e}\sqrt{(t-1) \cdot m}$ pages in $O(n^2m^{1.5})$ time.*

Thus, the upper bound of $O(\sqrt{m})$ on the sizes of the crossing and the disjoint sets, and the pagenumber can be obtained in polynomial time if the chromatic number of the underlying graph is constant and an optimal coloring of the graph is given. Employing the terminologies in [11, 12], Theorem 1.2 says that in $O(n^2m^{1.5})$ time, one can determine a layout of G so that both the stacknumber and at the same time the queuenumber are $O(\sqrt{t \cdot m})$, provided that the partite sets are given. We also remark that the known upper bounds for the chromatic number of G can be used for t in Theorem 1.2 (see Corollary 4.1) to establish the $O(\sqrt{m})$ upper bound, in polynomial time, for a large class of graphs containing all degree bounded and all genus bounded graphs.

Section 2 contains our probabilistic proofs. Section 3 contains our main result which is Theorem 1.1. The result is extended to arbitrary graphs in Section 4.

2 Probabilistic Proofs

Malitz [13] gave a probabilistic proof and a Las Vegas algorithm for embedding G in a book of $O(\sqrt{m})$ pages. It is not clear how to convert his probabilistic proof to a deterministic polynomial time algorithm. The main result of this section is a probabilistic proof to establish the $O(\sqrt{m})$ upper bound on the size of the largest crossing and disjoint sets, and the pagenumber, for a bipartite graph.

Let $G = (V_1, V_2, E)$ be a bipartite graph on the partite sets V_1 and V_2, and edge set E. We will denote the vertex set of G by V, that is, $V = V_1 \cup V_2$. For integers x and y, where $x \leq y$, let $[x, y]$ denote the set $\{x, x+1, \ldots, y\}$. We will assume throughout this paper that, $V_1 = [1, n_1]$ and $V_2 = [n_1 + 1, n]$, where $n = n_1 + n_2$. Let $V_1 \subseteq A \subseteq V$ and h be a linear ordering of A. We say that h is a *canonical* linear ordering on A, if $h(i) = i$ for any $i \in V_1$.

The following simple lemma is similar to Lemma 2.2 in [13] and exhibits the relationship between the linear drawings and the book embeddings.

Lemma 2.1 *Let $G = (V_1, V_2, E)$ and h be a canonical linear ordering on V. Let $M = \{a_i b_i \mid a_i \in V_1, b_i \in V_2, 1 \leq i \leq k\}$ be a k-matching in G, with $a_1 < a_2 < \ldots < a_k$.*

(i) M is a crossing set in (G, h) if and only if $h(b_1) < h(b_2) < \cdots < h(b_k)$.

(ii) M is a disjoint set in (G, h) if and only if $h(b_1) > h(b_2) > \cdots > h(b_k)$.

(iii) G can be embedded in a book of $C(G, h)$ pages in $O(m \log\log n)$ time.

Proof. Parts (i) and (ii) are easy to see. To verify (iii) for any two edge ab and cd, where $a, c \in V_1$, and $b, d \in V_2$, define $ab \succ cd$ if $a \leq c$ and $h(d) \leq h(b)$. Note that $\succ$ is a partial order on the edges of G in the drawing (G, h). In particular, observe that ab and cd are comparable if and only if ab does not cross cd in the drawing, that a chain with respect to $\succ$ can be embedded in one page, and that an anti-chain corresponds to a crossing set. (The fact that any edge pair can cross only once is crucial to this observation.) By Dilworth's Theorem [7], the minimum number of chains the edges can be partitioned to equals the size of the largest anti-chain, and thus, the claim on the number of pages follows. The time complexity can be achieved using a dual form of the algorithm by Heath and Rosenberg [12] (page 932). $\square$

In view of Part (iii) in Lemma 2.1, to solve the book embedding problem, we need to construct a linear drawing in which the size of the largest crossing set is small. The existence of a linear drawing (G, H) for which $C(G, H) \leq \lceil \mathbf{e}\sqrt{m} \rceil$ and $I(G, H) \leq \lceil \mathbf{e}\sqrt{m} \rceil$ can be shown by the following probabilistic argument. Let h be the identity function on V_1, that is, $h(i) = i$ for any $i \in [1, n_1]$. Now randomly and uniformly extend h to a linear ordering H of $V = [1, n]$. Thus, we pick any of the $n_2!$ possible extensions with probability $\frac{1}{n_2!}$. Then, using Part (i) in Lemma 2.1 it is easy to verify that the probability that any k-matching in G is a crossing set in (G, H) will be

$$\frac{\binom{n_2}{k}(n_2-k)!}{n_2!} = \frac{1}{k!}.$$

Since there are at most $\binom{m}{k}$ matchings of size k in G, the expected number of crossing sets of size k in a random linear drawing (G, H) is at most

$$\binom{m}{k}\frac{1}{k!} = \frac{m!}{k!(m-k)!k!} \leq \frac{m^k}{k!^2} \leq \frac{\mathbf{e}^{2k}}{2\pi k}(\frac{m}{k^2})^k,$$

where the last inequality is obtained using Stirling's formula,

$$k! \geq \left(\frac{k}{\mathbf{e}}\right)^k \sqrt{2\pi k}.$$

Note that for $k = \mathbf{e}\sqrt{m}$ the upper bound is $\frac{1}{2\pi\mathbf{e}} < \frac{1}{2}$. Similarly, using (ii) in Lemma 2.1 we can show that the expected number of disjoint sets of size k is also less than $\frac{1}{2}$, and using the linearity of the expectation, the sum is less than 1. Therefore, for $k = \lceil \mathbf{e}\sqrt{m} \rceil$, there exists a linear drawing with no crossing set and no disjoint set of size k. Note that the result holds for all values of m. Also note that by Part (iii) in Lemma 2.1 G can be embedded in a book of $\lceil \mathbf{e}\sqrt{m} \rceil$ pages.

The next result of this section is a more "computationally effective" form of the above arguments that can be used as a guideline for designing a deterministic algorithm. Before presenting the theorem, we need some definitions and a Lemma. Let $G = (V_1, V_2, E)$, and let $A \subseteq V_1$, and $B \subseteq V_2$. For an integer $k \geq 1$, define $M_k^{A,B}(G)$ to be the number of k-matchings in G composed of edges with one end point in A and the other in B. For any $A \subseteq V$, let $G[A]$ denote the induced subgraph of G on A.

For any k-matching M in G, $k \geq 1$, define the *right boundary* of M to be

$$\max_{uv \in M, u \in V_1} u,$$

and the *left boundary* of M to be

$$\min_{uv \in M, u \in V_1} u,$$

where the maximum and minimum are taken over all edges $uv \in M$ with $u \in V_1 = [1, n_1]$. Let $V_1 \subseteq A \subseteq V$ and h be a canonical linear ordering on A. The left (right) boundary of a crossing set or a disjoint set of size k in $(G[A], h)$ is defined to be the left (right) boundary of the k-matching that forms it. For integers $x \geq 1$ and $k \geq 1$, let $C_k^x(A, h)$ denote the number of crossing sets in the drawing $(G[A], h)$ that have size k and right boundary x. Similarly define $I_k^x(A, h)$ to be the number of disjoint sets in the drawing $(G[A], h)$ that have size k and left boundary x.

Let $V_1 \subseteq A \subseteq B \subseteq V$, h be a canonical linear ordering of A, and f be a canonical linear ordering of B. If $h(i) = f(i)$ for all $i \in A$, then we say f is an extension of h.

The following simple result regarding the structure of the crossing sets (twists) and the disjoint sets (rainbows) was noted by Heath and Rosenberg [12]. We will state it as a lemma and omit its proof.

Lemma 2.2 *Let $G = (V_1, V_2, E)$, A and B satisfy $V_1 \subseteq A \subseteq B \subseteq V$, h be a canonical linear ordering of A, and f be an extension of h to a canonical linear ordering on B. Let $k > t > 0$ and M_1 be a crossing set (disjoint set) of size $(k - t)$ in the drawing $(G[A], h)$. Let M_2 be a t-matching in $G[V_1 \cup (B - A)]$ that is a crossing set (disjoint set) of size t in the drawing $(G[B], f)$. Then $M_1 \cup M_2$ is a crossing set (disjoint set) of size k in the drawing $(G[B], f)$, if and only if, the left boundary of M_2 is greater than the right boundary of M_1 (right boundary of M_2 is less than the left boundary of M_1).*

□

We refere the reader to [9] for basic concepts on probability theory and to [3] for the probabilistic methods. For any random variable X, let $\mathbf{E}[X]$ denote the expected value of X. We can now state the main result of this section.

Theorem 2.1 *Let $G = (V_1, V_2, E)$, and let h be a canonical linear ordering of A, $V_1 \subseteq A \subseteq V$. Let H be a random extension of h to a canonical linear order of V which is obtained by randomly selecting any of the $(n - |A|)!$ extensions with the (uniform) probability of $1/(n - |A|)!$. Denote by $X_k(G, H)$ and by $Y_k(G, H)$ the random variables that count the number of crossing sets, and the number of disjoint sets, of size k, respectively, in the random linear drawing (G, H). Then*

$$\begin{aligned}\mathbf{E}[X_k(G,H) + Y_k(G,H)] &= \sum_{1\le x\le n_1} C_k^x(A,h) + \sum_{1\le x\le n_1} I_k^x(A,h) + 2\frac{M_k^{V_1,V-A}(G)}{k!} \\ &+ \sum_{\substack{1\le t\le k-1\\1\le x\le n_1}} \frac{C_{k-t}^x(A,h)\cdot M_t^{[x+1,n_1],V-A}(G)}{t!} \\ &+ \sum_{\substack{1\le t\le k-1\\1\le x\le n_1}} \frac{I_{k-t}^x(A,h)\cdot M_t^{[1,x-1],V-A}(G)}{t!}.\end{aligned} \tag{1}$$

Proof. By linearity of expectation, we have $\mathbf{E}[X_k(G,H) + Y_k(G,H)] = \mathbf{E}[X_k(G,H)] + \mathbf{E}[Y_k(G,H)]$. Note that $\mathbf{E}[X_k(G,H)]$ equals the number of the crossing sets of size k in $(G[A], h)$, plus the expected number of the crossing sets of size k generate by the matchings in $G[V_1 \cup (V - A)]$, plus the expected number of crossing sets of size k in (G, H) formed by the matchings in $G[A]$ and the matchings in $G[V_1 \cup (V - A)]$. Formally,

$$\mathbf{E}[X_k(G,H)] = \sum_{1\le x\le n_1} C_k^x(A,h) + \frac{M_k^{V_1,V-A}(G)}{k!} + \sum_{\substack{1\le t\le k-1\\1\le x\le n_1}} \frac{C_{k-t}^x(A,h)\cdot M_t^{[x+1,n_1],V-A}(G)}{t!}. \tag{2}$$

The first term in (2) is the number of the existing crossing sets of size k in (G, h). Next, note that for any k-matching in $G[V_1 \cup (V - A)]$, the probability that the k-matching is a crossing set of size k in (G, H) is

$$\frac{\binom{n-|A|}{k}(n-|A|-k)!}{(n-|A|)!} = \frac{1}{k!}.$$

Thus, the second term in (2) is the expected value for the number of crossing sets of size k in (G, H) that are generated by the k-matchings in $G[V_1 \cup (V - A)]$.

Finally, the third term in (2) is the expected value for the number of crossing sets of size k formed by an existing crossing set of size $k - t$ in (G, h) and a t-matching in $G[V_1 \cup (V - A)]$ that is a crossing set of size t in (G, H). To see this, let M_1 be a crossing set of size $k - t$ in $(G[A], h)$ whose right boundary is x. Let M_2 be a t-matching in $G[V_1 \cup (V - A)]$ and note that the probability that M_2 is a crossing set of size t in (G, H) is $\frac{1}{t!}$. Applying Lemma 2.2 with B being V, $M_1 \cup M_2$ is a crossing set of size k in (G, H) if and only if M_2 is a crossing set in (G, H) whose left boundary is at least $x + 1$.

Similarly, one can argue that

$$\mathbf{E}[Y_k(G, H)] = \sum_{1 \leq x \leq n_1} I_k^x(A, h) + \frac{M_k^{V_1, V-A}(G)}{k!} + \sum_{\substack{1 \leq t \leq k-1 \\ 1 \leq x \leq n_1}} \frac{I_{k-t}^x(A, h) \cdot M_t^{[1,x-1], V-A}(G)}{t!},$$

hence proving the Theorem. □

3 Derandomization

Note that if we take $A = V_1$ and $k = \lceil \mathrm{e}\sqrt{m} \rceil$, then the right hand side of (1) in Theorem 2.1 is less than 1. Therefore, Theorem 2.1 implies the previous probabilistic result. The main result of this section is to construct a suitable linear drawing using a deterministic polynomial time algorithm.

Applying the method of conditional probabilities [3] for removing the randomness in Theorem 2.1 requires counting the number of matchings of different size which is a #P-complete problem [23]. Nevertheless, we avoid the direct enumeration of matchings in the original graph to turn our probabilistic proof into a deterministic algorithm. More precisely, we convert the bipartite graph G to a forest G' that preserves certain properties of G, and then compute the conditional probabilities for G'.

For $G = (V_1, V_2, E)$ and any $i \in V_1$, let $d(i)$ denote the degree of i. Intuitively to construct G', we replace every vertex $i \in V_1$ with a set T_i of $d(i)$ vertices so that no two vertices in T_i are adjacent, and each vertex in T_i is adjacent to exactly one neighbor of i in V_2. Thus for any $i \in V_1$ the degree of any vertex in T_i is one.

Formally, for any $i \in [1, n_1]$ define T_i to be the set of all integers l, so that $\sum_{j=1}^{i-1} d(j) < l \leq \sum_{j=1}^{i} d(j)$, and denote by S_i the set $\{j + m - n_1 \mid ij \in E\}$. (Each vertex in S_i is a neighbor of i in G, which is renamed by being shifted $m - n_1$ units.) Note that, for any $i \in V_1$, $|T_i| = |S_i| = d(i)$. The bipartite graph $G = (V_1, V_2, E)$ has a *mate* that is a bipartite graph $G' = (V_1', V_2', E')$ on the vertex set $V' = V_1' \cup V_2'$, the edge set E' and is defined as follows. $V_1' = \cup_{i=1}^{n_1} T_i$ and $V_2' = \cup_{i=1}^{n_1} S_i$. Note that $V_1' = [1, m]$, $V_2' = [m+1, m+n_2]$, and $V' = [1, m+n_2]$. To construct E', for $i = 1, 2, \ldots, n_1$, we put a matching of size $d(i)$ between the vertices in T_i and the vertices in S_i. The choice of which vertex in T_i is matched to which vertex in S_i is arbitrary. Note that, $|E'| = \cup_{i=1}^{n_1} d(i) = m$ and that G' is a forest consisting of n_2 stars with their roots in V_2' and their leaves in V_1'. Also, observe that by collapsing all vertices in T_i to a single vertex, $i = 1, 2, \ldots, n_1$, we obtain the original graph G. Figure 1 shows a graph $G = K_{2,3}$ and its mate graph.

The following simple lemma explores some important properties of G'.

Lemma 3.1 *Let $G' = (V_1', V_2', E')$ be the mate of $G = (V_1, V_2, , E)$. Denote $V_1' \cup V_2'$ by V'. For any $W \subseteq V_1'$ and $i \in V_2'$, let $d_W(i)$ denote the number of vertices in W adjacent to i in G'.*

(i) Let h be a canonical linear ordering of V', then there is a canonical linear ordering f on V so that

$$C(G, f) \leq C(G', h) \text{ and } I(G, f) \leq I(G', h).$$

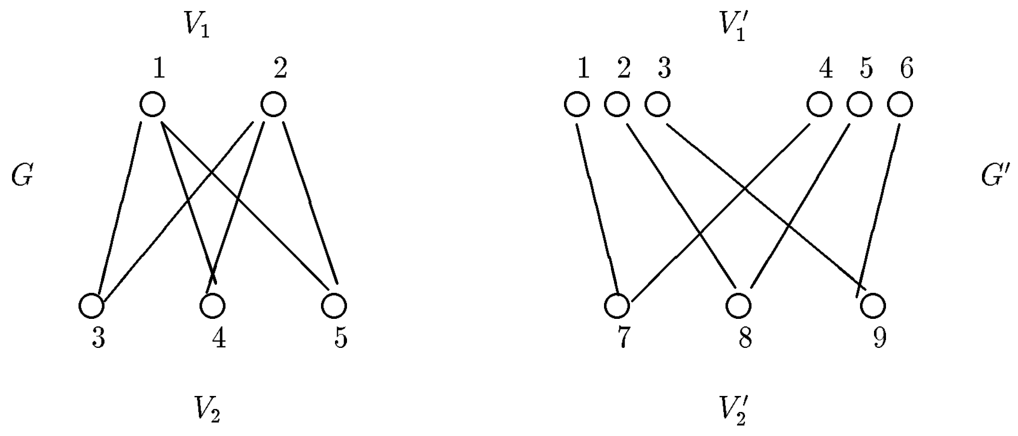

Figure 1: Graph $G = K_{2,3}$ and its mate graph G'.

(ii) Let $W \subseteq V_1'$, $B \subseteq V_2'$ and let $v \in B$, then,

$$M_k^{W,B-\{v\}}(G') = M_k^{W,B}(G') - d_W(v) \cdot M_{k-1}^{W,B-\{v\}}(G').$$

Proof. To prove (i) for $i = 1, 2, \ldots, n_1$, collapse all vertices in T_i in the linear drawing (G', h) to a single vertex i to obtain a linear drawing of a graph isomorphic to G and let f be the linear ordering of V which identifies this linear drawing. Proof of (ii) is easy and uses the fact that the degree of any vertex in V_1' is one. $\square$.

Let A satisfy $V_1' \subseteq A \subseteq V'$, let h be a canonical linear ordering on A, and let $v \in V' - A$. We denote by h_v the canonical linear ordering that is the extension of h from A to $A \cup \{v\}$. Note that for any $y \in A$, $h_v(y) = h(y)$, and that $h_v(v) = |A| + 1$.

Before presenting our main result, we prove another lemma which describes how to count the number of crossing and disjoint sets in the linear drawings of subgraphs of G'. The lemma will be used for bookkeeping in our algorithm.

Lemma 3.2 *Let $G' = (V_1', V_2', E')$ be the mate of $G = (V_1, V_2, E)$. Let h be a canonical linear ordering on A, $V_1' \subseteq A \subseteq V'$. Then the following hold.*

(i) For any $v \in V' - A$, and $x \in V_1'$

$$C_k^x(A \cup \{v\}, h_v) = \begin{cases} \sum_{y=1}^{x-1} C_{k-1}^y(A, h) & \textit{if } xv \in E', \\ C_k^x(A, h) & \textit{otherwise.} \end{cases}$$

(ii) For any $v \in V' - A$, and $x \in V_1'$

$$I_k^x(A \cup \{v\}, h_v) = \begin{cases} \sum_{y=x+1}^{m} I_{k-1}^y(A, h) & \textit{if } xv \in E', \\ I_k^x(A, h) & \textit{otherwise.} \end{cases}$$

Proof. To verify (i), first assume that $e = xv \in E'$. Then, the degree of x in $G'[A]$ is zero, since degree of x in G' is one. This means that there is no k-matching in $G'[A]$ and thus no crossing set of size k in $(G'[A], h)$ that has the right boundary x, and therefore $C_k^x(A, h) = 0$. Therefore, by Lemma 2.2, any crossing set of size k in $(G'[A \cup \{v\}], h_v)$ of right boundary x must be generated by a crossing set of size $k-1$ in $(G'[A], h)$ of right boundary y, $y \leq x-1$, together with edge e. Now assume that, $xv \notin E'$, and let $yv \in E'$, where $y \in V_1'$, $y \neq x$, is the unique vertex adjacent to v. We claim that edge e can not be an edge in a crossing set of size k and of right boundary x in the drawing $(G'[A \cup v], h_v)$. Assume to the contrary that $M_1 \cup \{e\}$ is a crossing set of size k in the drawing $(G'[A \cup v], h_v)$, and has right boundary x, where M_1 is a $(k-1)$-matching in $G'[A]$. It is then easily seen using Lemma 2.2 that, y must be greater than the right boundary of M_1. It follows that y must be the right boundary of $M_1 \cup \{e\}$ which is x. Consequently, $xv \in E'$ which is a contradiction. Proof of (ii) is similar to (i). □

We can now prove Theorem 1.1.

Proof of Theorem 1.1. In view of Part (iv) in Lemma 2.1, we only need to show suitable upper bounds on the size of largest crossing and disjoint sets. In view of Part (i) in Lemma 3.1, it suffices to verify the claim for G'. Our algorithm essentially applies Theorem 2.1 to G' in $O(n^2 m^{1.5})$ time. The algorithm has n_2 iterations. At iteration i, $1 \leq i \leq n_2$, the algorithm constructs a set A_i, $V_1' \subseteq A_i \subseteq V'$, $|A_i| = n_1 + i$, and a canonical linear ordering

h^i on A_i. When the algorithm terminates, the linear ordering $h = h^{n_2}$ is the desired linear ordering. Initially we maintain that $A_0 = V_1'$, $k = \lceil \mathbf{e}\sqrt{m} \rceil$ and h^0 to be the identity function on V_1'. We denote by H^0 the random extension of h^0 to a linear ordering of V'. We also denote by Z^0 the expected number of crossing sets plus the number of disjoint sets of size k, in the linear drawing (G', H^0) and note that $Z^0 < 1$ by Theorem 2.1.

At iteration i, $1 \le i \le n_2$, we construct $A_i = A_{i-1} \cup \{v_i\}$ and $h^i = h^{i-1}_{v_i}$, where v_i is a particular vertex in $V' - A_{i-1}$. Employing the notations in Theorem 2.1, for any $v \in V' - A_i$, $i \ge 0$, let H^i_v denote the random extension of h^i_v to a linear ordering of V'. To select vertex v_i at iteration $i \ge 1$, for any $v \in V' - A_{i-1}$, we compute using Theorem 2.1

$$Z^i_v = \mathbf{E}[X_k(G', H^{i-1}_v) + Y_k(G', H^{i-1}_v)],$$

and note that Z^i_v is the expected number of crossing sets plus disjoint sets of size k, provided that we randomly and uniformly extend h^{i-1}_v to a linear ordering of V'. Let Z^i be the minimum of all Z^i_v's, $v \in V_2' - A_{i-1}$; we select v_i to be the vertex that achieves the minimum. By the expansion of conditional probability,

$$Z^0 \ge Z^1 \ge Z^2 \ge \cdots \ge Z^{n_2}.$$

Moreover Z^{n_2} is the number of crossing sets of size k plus the number of disjoint sets of size k in (G', h), where $h = h^{n_2}$ is the final linear ordering . Consequently, $Z^{n_2} \le Z^0 < 1$, implying that (G', h) does not have any crossing set or any disjoint set of size k.

To finish the proof it suffices to show that at any iteration i and for any $v \in V' - A_{i-1}$, Z^i_v is computed in $O(mk)$ time. Thus the total time at iteration i is $O(nmk)$, and the algorithm has worst-case time complexity $O(n^2m^{1.5})$.

We store $C^x_t(A_i, h^i)$ in a table with entries for $x = 1, 2, \ldots, m$ and $t = 1, 2, \ldots, k$. In addition, we store the partial sum $\sum_{y=1}^{x} C^y_t(A_i, h^i)$ in another table, for $x = 1, 2, \ldots, m$ and $t = 1, 2, \ldots, k$. This way, using Part (i) in Lemma 3.2, at iteration i, for every x, t and v,

$C_t^x(A_i \cup \{v\}, h_v^i)$ can be computed in $O(1)$ time. Similarly using two additional tables, we can compute $I_t^x(A_i \cup \{v\}, h_v^i)$, in $O(1)$ time for every x, t, v at iteration i.

We also store $M_t^{[x,m],V'-A_i}(G')$ in a table with entries for $x = 1, 2, \ldots, m$ and $t = 1, 2, \ldots, k$. From Part (ii) of Lemma 3.1, if we compute the entries in increasing t value, then for every x, t, v, $M_t^{[x,m],V'-A_i-\{v\}}(G')$ can be computed in $O(1)$ time.

Therefore, at iteration i, for any $v \in V' - A_{i-1}$, we can compute Z_v^i using Theorem 2.1 in time $O(mk)$ as claimed. At the end of each iteration, we update the tables in time $O(mk)$. Hence the worst-case time complexity for the algorithm is $O(n^2mk)$ or $O(n^2m^{1.5})$.

□

4 Extension To Arbitrary Graphs

In this section we extend our results for bipartite graphs to arbitrary graphs.

Proof of Theorem 1.2. Assume that $V = V_1 \cup V_2 \cup \cdots \cup V_t$, where each V_i, $1 \le i \le t$, is an independent set of vertices. Let E_i be the set of edges between $V_1 \cup \cdots \cup V_i$ and V_{i+1}, $i = 1, \ldots, t-1$. Start with the identity function h_0 on V_1 and apply the deterministic algorithm in Theorem 1.1 to the bipartite graph $G_1 = (V_1, V_2, E_1)$ to extend h_0 to a linear ordering h_1 on $V_1 \cup V_2$, then extend h_1 to a linear ordering h_2 on $V_1 \cup V_2 \cup V_3$ by applying Theorem 1.1 to the bipartite graph $G_2 = (V_1 \cup V_2, V_3, E_2)$, and so on. When we finish, we have a linear ordering $h = h_{t-1}$ of V and a linear drawing such that

$$C(G,h) \le \mathbf{e}\sqrt{|E_1|} + \mathbf{e}\sqrt{|E_2|} + \cdots + \mathbf{e}\sqrt{|E_{t-1}|} \le \mathbf{e}\sqrt{(t-1)|E|}.$$

Similarly, we obtain an inequality for $I(G,h)$. Moreover, the claim regarding the number of pages follows by observing that the application of Theorem 1.1 gives an embedding of G_i, $1 \le i \le t-1$, in $\mathbf{e}\sqrt{|E_i|}$ pages. It is easy to verify that the time complexity is $O(n^2m^{1.5})$. □

Let $\delta(G)$ denote the minimum degree of $G = (V, E)$, and define the *linkage of* G to be

$$\delta^*(G) = \max_{H \subseteq G} \{\delta(H)\},$$

where the maximum is taken over all subgraphs H of G (see Szekeres and Wilf [22]). Let $\chi(G)$ denote the chromatic number of G. It is known that [22] $\chi(G) \leq \delta^*(G) + 1$. Moreover, efficient algorithms for coloring G with $\delta^*(G)+1$ colors are known. See [15] for a linear time algorithm. Substituting the upper bound of $\delta^*(G) + 1$ for t in Theorem 1.2, we obtain:

Corollary 4.1 *Let $G = (V, E)$. There is an $O(n^2 m^{1.5})$ time algorithm that finds a linear drawing (G, h), so that $C(G, h) \leq \mathbf{e}\sqrt{\delta^*(G) \cdot m}$ and $I(G, h) \leq \mathbf{e}\sqrt{\delta^*(G) \cdot m}$. Moreover, G can be embedded in a book of $\mathbf{e}\sqrt{\delta^*(G) \cdot m}$ pages in $O(n^2 m^{1.5})$ time.*

□

Thus, for the class of graphs with bounded linkage which contains all bounded degree graphs and all genus bounded graphs, the upper bound of $O(\sqrt{m})$ can be achieved.

Acknowledgment. The authors thank the referees for useful suggestions, especially for pointing out an error in an earlier version of the paper, and for improving the time complexity in Part (iii) of Lemma 2.1.

References

[1] P. K. Agarwal, B. Arnov, J. Pach, R. Pollack, and M. Sharir, Quasi-planar graphs have a linear number of edges, *Combinatorica* **17** (1997), 1–9.

[2] N. Alon and P. Erdös, Disjoint edges in geometric graphs, *Disc. Comput. Geom.* **4** (1989), 287–290.

[3] N. Alon, J. H. Spencer and P. Erdős, *The Probabilistic Method*, Wiley and Sons, New York, 1992.

[4] F. Bernhart and B. Kainen, The book thickness of a graph, *J. Combin. Theory* **B27** (1979), 320–331.

[5] V. Capoyleas and J. Pach, A Turan type theorem on chords of a convex polygon, *J. Combin Theory* **B56** (1992), 9–15.

[6] F. R. K. Chung, F. T. Leighton and A. Rosenberg, Embedding graphs in books: A layout problem with applications to VLSI design, *SIAM J. Alg. Disc. Meth.* **8** (1987), 33–58.

[7] R. Dilworth, A decomposition theorem for partially ordered sets. *Ann. Math. Ser. 2* **51** (1950), 26–32.

[8] M. R. Garey, D. S. Johnson, G. L. Miller and C. H. Papadimitriou, The complexity of coloring circular arc graphs and chords, *SIAM J. Alg. Disc. Meth.* **1** (1980), 33–58.

[9] W. Feller, *An Introduction to Probability Theory and Its Applications*, Wiley and Sons, New York, 1968.

[10] L. S. Heath and S. Istrail, The pagenumber of genus g graph is $O(g)$, *JACM* **39** (1992), 479–500.

[11] L. S. Health, F. T. Leighton, and A. L. Rosenberg, Comparing queues and stacks as mechanisms for laying out graphs, *SIAM J. Disc. Math.* **5** (1992), 398–412.

[12] L. S. Heath and A. L. Rosenberg, Laying out graphs using queues,, *SIAM J. Comput.* **21** (1992), 927–958.

[13] S. M. Malitz, On the pagenumber of graphs, *J. Algorithms* **17** (1) (1994), 71–84.

[14] S. M. Malitz, Genus g graphs have pagenumber $O(\sqrt{g})$, *J. Algorithms* **17** (1) (1994), 85–109.

[15] D. W. Matula and L. L. Beck, Smallest-last ordering and clustering and graph coloring algorithms, *JACM* **30** (1983), 417–427.

[16] J. Pach, P. K. Agarwal, Combinatorial Geometry, *Wiley and Sons Inc., New York, N.Y.*, 1995.

[17] J. Pach, F. Shahrokhi and M. Szegedy, Applications of crossing numbers, *Algorithmica* **16** (1996), 111–117.

[18] J. Pach and J. Töröcsik, Some geometric applications of Dilworth's theorem, *Disc. Comput. Geometry* **21** (1994), 1–7.

[19] P. Raghavan, Probabilistic construction of deterministic algorithms: Approximating packing integer programs, *J. Comput. Sys. Sci.* **37** (1988), 130–143.

[20] A. L. Rosenberg, The Diogenes approach to fault-tolerant arrays of processors, *IEEE Trans. Computers* **C-32** (1983), 902–910.

[21] P. Rosenstiehl and R. E. Tarjan, Gauss codes, planar Hamiltonian graphs and stack-sortable permutations, *J. Algorithms* **5** (1984), 375–390.

[22] G. Szekeres and H. S. Wilf, An inequality for the chromatic number of a graph, *J. Comb. Theory* **4** (1968), 1–3.

[23] L. Valiant, On the complexity of counting the matching, *Theoretical Comput. Sci.* **8** (1979), 189–201.

[24] M. Yannakakis, Embedding planar graphs in four pages, *J. Comput. Sys. Sci.* **38** (1989), 36–67.

[25] P. Valtr, On geometric graphs with no k pairwise parallel edges, *Disc. Comput. Geometry* **19** (1998), 461–469.